

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:42:31 ; Search time 161.856 Seconds
(without alignments)
4214.186 Million cell updates/sec

Title: US-09-982-091A-4
Perfect score: 6836
Sequence: 1 MTGEVGSSEVHLEINDPNVIS.....LTWSGAPIPGFFRLSPDPHG 1332

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6836	100.0	1332	2	Q9HAW4	Q9HAW4 homo sapien
2	4838	70.8	1315	2	Q80YR7	Q80YR7 mus musculu
3	3512	51.4	693	2	Q8IWI1	Q8IWI1 homo sapien
4	2687	39.3	1285	2	Q9DF50	Q9DF50 xenopus lae
5	2340	34.2	462	2	Q9BX02	Q9BX02 homo sapien
6	1618.5	23.7	411	2	Q69GM2	Q69GM2 mus musculu
7	1500	21.9	289	2	Q9BX01	Q9BX01 homo sapien
8	836	12.2	165	2	Q9BX03	Q9BX03 homo sapien
9	784	11.5	270	2	Q8BPX6	Q8BPX6 mus musculu
10	700	10.2	1465	2	Q8IRB5	Q8IRB5 drosophila
11	534.5	7.8	180	2	Q8BRD4	Q8BRD4 mus musculu
12	526	7.7	835	2	Q7PR25	Q7PR25 anopheles g
13	497	7.3	105	2	Q6P6H5	Q6P6H5 homo sapien
14	487.5	7.1	7210	2	Q9V7G8	Q9V7G8 drosophila
15	487.5	7.1	9270	2	Q8MLD9	Q8MLD9 drosophila
16	454	6.6	2081	2	Q9LH98	Q9LH98 arabidopsis
17	448.5	6.6	725	2	Q17838	Q17838 caenorhabdi
18	448	6.6	701	2	Q8MYR6	Q8MYR6 drosophila
19	446.5	6.5	3404	2	Q7RJB3	Q7RJB3 plasmodium
20	419	6.1	1464	2	Q871F7	Q871F7 neurospora
21	415	6.1	1661	2	Q06166	Q06166 plasmodium
22	413.5	6.0	10578	2	Q8ISF5	Q8ISF5 caenorhabdi
23	413.5	6.0	18519	2	Q8ISF6	Q8ISF6 caenorhabdi
24	413.5	6.0	18534	2	Q8ISF7	Q8ISF7 caenorhabdi
25	411	6.0	2301	2	Q6BG00	Q6BG00 paramedum
26	410	6.0	1451	2	Q812D8	Q812D8 plasmodium
27	409	6.0	5507	2	Q8IHN3	Q8IHN3 plasmodium
28	407.5	6.0	5458	2	Q9U459	Q9U459 plasmodium
29	402.5	5.9	1481	2	Q8EVB9	Q8EVB9 mycoplasma
30	396.5	5.8	1510	2	Q25920	Q25920 plasmodium
31	395.5	5.8	1871	2	Q9SRD5	Q9SRD5 arabidopsis

32	395.5	5.8	2760	2	Q8ISY2	Q8ISY2 plasmodium
33	394	5.8	5412	1	FUTS DROME	Q9W596 drosophila
34	391	5.7	791	2	Q9DGL1	Q9DGL1 fungu rubrip
35	391	5.7	1434	2	Q8I492	Q8I492 plasmodium
36	388.5	5.7	2468	1	MAPB HUMAN	P46821 homo sapien
37	385	5.6	3504	2	Q8IL45	Q8IL45 plasmodium
38	384.5	5.6	2116	1	MYS2 DICDI	P08799 dictyosteli
39	383.5	5.6	1431	2	Q96T23	Q96T23 homo sapien
40	381	5.6	2464	2	Q64ID6	Q64ID6 macropus eu
41	381	5.6	2910	2	Q9FNDS	Q9FNDS arabidopsis
42	380	5.6	1989	2	Q8I4U7	Q8I4U7 plasmodium
43	380	5.6	2095	2	Q7RN75	Q7RN75 plasmodium
44	376	5.5	1805	1	NEST RAT	P21263 rattus norv
45	375	5.5	2459	1	MAPB RAT	P15205 rattus norv

ALIGNMENTS

RESULT 1									
ID	Q9HAW4	PRELIMINARY;	PRT;	1332	AA.				
AC	Q9HAW4;								
DT	01-MAR-2001	(TReMBUrel. 16, Created)							
DT	01-MAR-2001	(TReMBUrel. 16, Last sequence update)							
DT	01-OCT-2003	(TReMBUrel. 25, Last annotation update)							
DE	Hu-Claspin.								
OS	Homo sapiens (Human).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
OX	NCBI_TaxID=9606;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=21000493; PubMed=11090622;								
RA	Kumagai A., Dunphy W.G.;								
RT	"Claspin, a novel protein required for the activation of Chk1 during a								
RT	DNA replication checkpoint response in Xenopus egg extracts.";								
RL	Mol. Cell 6:839-849(2000).								
DR	EMBL; AF297866; AAG24515.1; -.								
DR	Genew; HGNC:19715; CLSPN.								
DR	GO; GO:0005515; F:protein binding; NAS.								
DR	SEQUENCE 1332 AA; 150174 MW; 4EA38DC16A3936C3 CRC64;								
Query Match 100.0%; Score 6836; DB 2; Length 1332;									
Best Local Similarity 100.0%; Pred. No. 8.3e-227;									
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MTGEVGSEVHLEINDPNVISQEEADSPSDSGQSYETIGPLSEGDSEEIFVSKKLKNRK	60						
Db	1	MTGEVGSEVHLEINDPNVISQEEADSPSDSGQSYETIGPLSEGDSEEIFVSKKLKNRK	60						
QY	61	VLQSDSEETEDTNASPEKTTYDSAEENKENLYAGKNTKIRIYKTVADSDSEYMEKSLY	120						
Db	61	VLQSDSEETEDTNASPEKTTYDSAEENKENLYAGKNTKIRIYKTVADSDSEYMEKSLY	120						
QY	121	QENLEAOVKPCLELSLQSGNSTFTTDRKSSKKHIHDKGTAGKAKVYSKRLKEEKRM	180						
Db	121	QENLEAOVKPCLELSLQSGNSTFTTDRKSSKKHIHDKGTAGKAKVYSKRLKEEKRM	180						
QY	181	EKIRQLKKKETKNQEDDVEQPFNDGCLLVDKDLFETGLEDENNSPLEDEESLESIRAAV	240						
Db	181	EKIRQLKKKETKNQEDDVEQPFNDGCLLVDKDLFETGLEDENNSPLEDEESLESIRAAV	240						
QY	241	KNKVKKKKKKPSLESVHSFEEGSELSKGTTRKERKARLSKEALKQLHSETORLRES	300						
Db	241	KNKVKKKKKKPSLESVHSFEEGSELSKGTTRKERKARLSKEALKQLHSETORLRES	300						
QY	301	ALNLPYHMPENKTIHDFFKRKRPPTCHGNAMALLKSKYQSSHHKEIIDTANTTEMNSDH	360						
Db	301	ALNLPYHMPENKTIHDFFKRKRPPTCHGNAMALLKSKYQSSHHKEIIDTANTTEMNSDH	360						
QY	361	HSKGSSEOTTGAENEVETNALPVVSKETQIITGSDSECRDLVKNLELEIOEKOKOSDIRP	420						

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:27:19 ; Search time 34.1016 Seconds
(without alignments)
2915.770 Million cell updates/sec

Title: US-09-982-091a-4
Perfect score: 6836
Sequence: 1 MTGEVGEVHLEINDPNVIS.....LTWGCAPIPGFFRLSFDPHG 1332

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5555	81.3	1089	4 US-09-949-016-10326	Sequence 10326, A
2	388.5	5.7	2468	4 US-09-976-594-726	Sequence 726, App
3	388.5	5.7	2468	4 US-09-538-092-1135	Sequence 1135, Ap
4	388.5	5.7	2522	4 US-09-949-016-10237	Sequence 10237, A
5	382.5	5.6	1400	4 US-09-764-176-7	Sequence 7, Appl1
6	376	5.5	1805	1 US-07-853-913-2	Sequence 2, Appl1
7	352.5	5.2	2375	4 US-09-538-092-1131	Sequence 1131, Ap
8	349	5.1	1898	1 US-08-056-200-94	Sequence 94, Appl
9	349	5.1	1898	2 US-08-800-644-94	Sequence 94, Appl
10	349	5.1	1898	4 US-09-538-092-1280	Sequence 1280, Ap
11	348	5.1	3259	4 US-09-949-016-6507	Sequence 6507, Ap
12	341	5.0	2733	4 US-09-949-016-11433	Sequence 11433, A
13	333.5	4.9	3878	4 US-09-914-259-11	Sequence 11, Appl
14	329.5	4.8	2476	4 US-09-824-574-7	Sequence 7, Appl1
15	327	4.8	2954	4 US-09-150-867-1	Sequence 1, Appl1
16	323.5	4.7	8991	4 US-08-714-741-32	Sequence 32, Appl
17	313.5	4.6	1939	3 US-09-310-187A-1	Sequence 1, Appl1
18	312.5	4.6	1935	4 US-09-538-092-916	Sequence 916, App
19	312	4.6	1786	3 US-08-973-462-8	Sequence 8, Appl1
20	311.5	4.6	1944	4 US-09-949-016-10929	Sequence 10929, A
21	310.5	4.5	1939	4 US-09-538-092-917	Sequence 917, App
22	308	4.5	2663	4 US-09-538-092-1252	Sequence 1252, App
23	304.5	4.5	1886	3 US-08-938-105-3	Sequence 3, Appl1
24	302	4.4	3696	3 US-09-134-001C-5080	Sequence 5080, Ap
25	301	4.4	1857	4 US-09-917-254-91	Sequence 91, Appl
26	301	4.4	1972	4 US-09-538-092-1084	Sequence 1084, Ap
27	301	4.4	1976	4 US-09-538-092-1078	Sequence 1078, Ap

28	301	4.4	1984	4 US-09-949-016-7111	Sequence 7111, Ap
29	301	4.4	1984	4 US-09-949-016-7112	Sequence 7112, Ap
30	301	4.4	1984	4 US-09-949-016-7113	Sequence 7113, Ap
31	300.5	4.4	1427	4 US-09-538-092-1044	Sequence 1044, Ap
32	300	4.4	3924	4 US-09-538-092-1246	Sequence 1246, Ap
33	299.5	4.4	1942	4 US-09-949-016-8135	Sequence 8135, Ap
34	299.5	4.4	2662	4 US-09-595-684B-31	Sequence 31, Appl
35	298	4.4	1618	1 US-07-853-913-4	Sequence 4, Appl1
36	298	4.4	1618	4 US-09-538-092-1143	Sequence 1143, Ap
37	297.5	4.4	1972	4 US-08-875-435B-3	Sequence 3, Appl1
38	297	4.3	1972	4 US-09-418-710-21	Sequence 21, Appl
39	297	4.3	1972	4 US-09-839-479-21	Sequence 21, Appl
40	296.5	4.3	1969	4 US-09-418-710-72	Sequence 72, Appl
41	296.5	4.3	1969	4 US-09-839-479-71	Sequence 71, Appl
42	296	4.3	1085	1 US-08-431-080-28	Sequence 28, Appl
43	296	4.3	1085	2 US-08-938-534-28	Sequence 28, Appl
44	296	4.3	1085	3 US-09-345-294-28	Sequence 28, Appl
45	296	4.3	2349	4 US-09-538-092-914	Sequence 914, Appl

ALIGNMENTS

```
RESULT 1
US-09-949-016-10326
; Sequence 10326, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241, 755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237, 768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231, 498
; PRIOR FILING DATE: 2000-09-08
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 10326
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10326 :
Query Match      81.3%; Score 5555; DB 4; Length 1089;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1081; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      248 KKKEPSLESQVHSFEEGSELSKGTTRKERKAARLSKEALKOLHSETORLIREALNLPYH 307
        :|||||
Db      5 OKKEPSLESQVHSFEEGSELSKGTTRKERKAARLSKEALKOLHSETORLIREALNLPYH 64

QY      308 MPENKTIHDFEKKRPPTCHGNAMALLKSSKYQSSHHEKITDTANTTEMNSDHSKGSSEQ 367
        :|||||
Db      65 MPENKTIHDFEKKRPPTCHGNAMALLKSSKYQSSHHEKITDTANTTEMNSDHSKGSSEQ 124

QY      368 TTGAENEVERTNALPVSKETQITTSDESCRKDLVKNEELEIOEKOKOSDIRPSPGDSV 427
        :|||||
Db      125 TTGAENEVERTNALPVSKETQITTSDESCRKDLVKNEELEIOEKOKOSDIRPSPGDSV 184

QY      428 LQOESNPLGNHSEECQVGLVAREPHALEGEGPQNPETDEKVEEPEQONKSSAVGPPE 487
        :|||||
Db      185 LQOESNPLGNHSEECQVGLVAREPHALEGEGPQNPETDEKVEEPEQONKSSAVGPPE 244

QY      488 KVRRTFLDRKQLGVDVSIKPRLGADSDSFVILEPETNRELEALKORFWKHANPAKPRRA 547
        :|||||
Db      245 KVRRTFLDRKQLGVDVSIKPRLGADSDSFVILEPETNRELEALKORFWKHANPAKPRRA 304

QY      548 GQTVNVNVIKMDGTGKELKADVVPVTLAPKLDGASHTKPGKELQVLKAKIQEAMKL 607
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:37:55 ; Search time 146.586 Seconds
(without alignments)
3793.838 Million cell updates/sec

Title: US-09-982-091A-4

Perfect score: 6836
Sequence: 1 MTGEVGSVHLEINDPNVIS.....LTWSGAPIPGFRLSFDPHG 1332

Scoring table: BLOSUM62
Gapop 10.0 , Gapexc 0.5

Searched: 1862994 seqs, 417510619 residues

Total number of hits satisfying chosen parameters: 1862994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6836	100.0	1332	US-09-982-091A-4	Sequence 4, Appl1
2	2687	39.3	1285	US-09-982-091A-2	Sequence 2, Appl1
3	850	12.4	169	US-10-106-698-5533	Sequence 5533, Ap
4	697	10.2	1812	US-11-097-143-858	Sequence 858, App
5	388.5	5.7	2468	US-10-755-889-615	Sequence 615, App
6	388.5	5.7	2468	US-10-489-740-216	Sequence 216, App
7	388.5	5.7	2519	US-10-450-763-46995	Sequence 46995, A
8	382.5	5.6	1398	US-10-094-466-30	Sequence 30, Appl
9	382.5	5.6	1400	US-09-764-176-7	Sequence 7, Appl1
10	375	5.3	2364	US-10-205-331-66	Sequence 66, Appl
11	365.5	5.3	1909	US-10-732-923-3341	Sequence 3341, Ap

12	359.5	5.3	1679	15	US-10-369-493-22080	Sequence 22080, A
13	358.5	5.2	1234	17	US-10-741-849-7132	Sequence 7132, Ap
14	354	5.2	2768	20	US-11-097-143-31983	Sequence 31983, A
15	353.5	5.2	2288	17	US-10-732-923-8885	Sequence 8885, Ap
16	353.5	5.2	2288	18	US-10-732-923-8885	Sequence 8885, Ap
17	353.5	5.2	2337	17	US-10-732-923-8888	Sequence 8888, Ap
18	353.5	5.2	2375	17	US-10-732-923-8880	Sequence 8880, Ap
19	353.5	5.2	2454	17	US-10-732-923-8884	Sequence 8884, Ap
20	353.5	5.2	2492	17	US-10-732-923-8882	Sequence 8882, Ap
21	353.5	5.2	2492	17	US-10-732-923-8882	Sequence 8882, Ap
22	353.5	5.2	2492	18	US-10-934-998-76	Sequence 76, Appl
23	352.5	5.2	2375	17	US-10-732-923-8881	Sequence 8881, Ap
24	352.5	5.2	2375	17	US-10-732-923-8881	Sequence 8881, Ap
25	352.5	5.2	2492	17	US-10-732-923-8883	Sequence 8883, Ap
26	352	5.1	3257	20	US-11-097-143-29298	Sequence 29298, A
27	351	5.1	2274	20	US-11-097-143-2763	Sequence 2763, Ap
28	350	5.1	1407	17	US-10-732-923-3359	Sequence 3359, Ap
29	348	5.1	2897	20	US-11-097-143-2334	Sequence 2334, Ap
30	347.5	5.1	2910	17	US-10-732-923-3342	Sequence 3342, Ap
31	347	5.1	1881	14	US-10-032-585-7646	Sequence 7646, Ap
32	346.5	5.1	2062	15	US-10-052-648A-52	Sequence 52, Appl
33	346	5.1	3111	20	US-11-097-143-7773	Sequence 7773, Ap
34	343.5	5.0	3907	14	US-10-171-311-2	Sequence 2, Appl1
35	343.5	5.0	3908	18	US-10-756-149-5739	Sequence 5739, Ap
36	343.5	5.0	3925	14	US-10-171-311-6	Sequence 6, Appl1
37	342	5.0	3899	14	US-10-171-311-4	Sequence 4, Appl1
38	342	5.0	3917	14	US-10-171-311-8	Sequence 8, Appl1
39	341	5.0	1183	20	US-11-097-143-3099	Sequence 3099, Ap
40	341	5.0	3225	16	US-10-408-765A-254	Sequence 254, App
41	337.5	4.9	1939	17	US-10-732-923-3340	Sequence 3340, Ap
42	333.5	4.9	3878	14	US-10-080-608A-11	Sequence 11, Appl
43	333.5	4.9	3911	15	US-10-370-685-100	Sequence 100, App
44	333.5	4.9	3911	16	US-10-408-765A-1839	Sequence 1839, Ap
45	333	4.9	1015	17	US-10-732-923-3329	Sequence 3329, Ap

ALIGNMENTS

RESULT 1
US-09-982-091A-4
Sequence 4, Application US/09982091A
Patent No. US20020151030A1
GENERAL INFORMATION:
APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
APPLICANT: KUMAGAI, Akiyo
APPLICANT: DUNPHY, William
TITLE OF INVENTION: CLASPIN PROTEINS AND METHODS OF USE THEREOF
FILE REFERENCE: CIT1320-1
CURRENT APPLICATION NUMBER: US/09/982, 091A
CURRENT FILING DATE: 2002-10-17
PRIOR APPLICATION NUMBER: US 60/241,246
PRIOR FILING DATE: 2000-10-17
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentin version 3.1
SEQ ID NO 4
LENGTH: 1332
TYPE: PRT
ORGANISM: Homo sapiens
US-09-982-091A-4

Query Match	100.0%	Score 6836,	DB 9,	Length 1332;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1332;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MTGEVGSVHLEINDPNVISQEEADSPDSGQGSYETIGPISGDSDEIFVSKLKNRK	60	
Db	1	MTGEVGSVHLEINDPNVISQEEADSPDSGQGSYETIGPISGDSDEIFVSKLKNRK	60	
QY	61	VLQDSSETEDTNNAPEKTTYDSAEENKENLYAGKNTKIKRIYKTVADSDSYMEKSLY	120	
Db	61	VLQDSSETEDTNNAPEKTTYDSAEENKENLYAGKNTKIKRIYKTVADSDSYMEKSLY	120	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:40:50 ; Search time 143.023 Seconds
(without alignments)
3601.964 Million cell updates/sec

Title: US-09-982-091A-4
Perfect score: 6836
Sequence: 1 MTGEVGSSEVHLEINDPNVIS.....LTWSGAPIPGFFRLSPDPHG 1332

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6836	100.0	1332	5 AAU97587	Aau97587 Human C1a
2	6812	99.6	1332	7 ADC31665	Adc31665 Human nov
3	2687	39.3	1285	5 AAU97586	Aau97586 Xenopus C
4	850	12.4	169	4 AAG74759	Aag74759 Human col
5	780	11.4	191	4 ABB99891	Abb99891 Human mac
6	697	10.2	1812	4 ABB58022	Abb58022 Drosophi1
7	388.5	5.7	2468	6 ABR64281	Abf64281 Angiogene
8	388.5	5.7	2468	7 ADE62723	Adc62723 Human Pro
9	388.5	5.7	2468	7 ADE62719	Adc62719 Human Pro
10	388.5	5.7	2468	7 ADE62727	Adc62727 Human Pro
11	388.5	5.7	2468	7 ADE62715	Adc62715 Human Pro
12	388.5	5.7	2468	8 ADL12997	Adl12997 Human ste
13	388.5	5.7	2468	8 ADN05260	Adn05260 Antipsori
14	388.5	5.7	2468	8 ADL14614	Adl14614 Human NF-
15	388.5	5.7	2519	4 ABG16636	Abg16636 Novel hum
16	388.5	5.7	2527	8 ADN04561	Adn04561 Antipsori
17	382.5	5.6	1398	5 ABG97491	Abg97491 Human NOV
18	382.5	5.6	1400	4 AAB83348	Aab83348 AAP-2 pro
19	376	5.5	1805	2 AAR27204	Aar27204 Rat nesti
20	376	5.5	1805	2 AAR60126	Aar60126 Rat nesti
21	376	5.5	1805	7 ADD48689	Add48689 Rat Prote
22	375	5.5	2364	6 ABM04816	Abm04816 Rat micro
23	375	5.5	2459	7 ADE62725	Adc62725 Rat Prote
24	375	5.5	2459	7 ADE62713	Adc62713 Rat Prote
25	375	5.5	2459	7 ADE62717	Adc62717 Rat Prote

26	375	5.5	2459	7 ADD48462	Add48462 Rat Prote
27	375	5.5	2459	7 ADE62721	Adc62721 Rat Prote
28	374.5	5.5	1445	6 ABR41365	Abf41365 Human DIT
29	359.5	5.3	1679	8 ADS43650	Ads43650 Bacterial
30	358.5	5.2	1234	8 ADP98957	Adp98957 C. albica
31	354	5.2	2768	4 ABB68397	Abb68397 Drosophi1
32	353.5	5.2	2288	8 ADK60224	Adk60224 Angiogene
33	353.5	5.2	2288	8 ADK60525	Adk60525 Angiogene
34	353.5	5.2	2288	8 ADP73148	Adp73148 Angiogene
35	353.5	5.2	2492	8 ADK60200	Adk60200 Angiogene
36	353.5	5.2	2492	8 ADK60501	Adk60501 Angiogene
37	353.5	5.2	2492	8 ADP73124	Adp73124 Angiogene
38	353.5	5.2	2492	8 ADQ89534	Adq89534 Human ATR
39	352.5	5.2	2375	7 ADJ68471	Adj68471 Human hea
40	352.5	5.2	3818	8 ADF45525	Adf45525 Human AKA
41	352	5.1	3257	4 ABB67502	Abb67502 Drosophi1
42	351	5.1	2274	4 ABB58657	Abb58657 Drosophi1
43	349	5.1	1898	2 AAY30795	Aay30795 A human t
44	349	5.1	1898	7 ADD48869	Add48869 Human Pro
45	348	5.1	2897	4 ABB58514	Abb58514 Drosophi1

ALIGNMENTS

RESULT 1	
ID	AAU97587
AC	AAU97587 standard; protein; 1332 AA.
DT	27-AUG-2002 (first entry)
DE	Human Claspin protein.
KW	Chk1 protein; SQ/TQ motif; isoelectric point; cell cycle progression; nuclear localisation signal; DNA replication checkpoint; benign neoplasm
KW	cell proliferative disorder; malignant neoplasm; human; claspin.
OS	Homo sapiens.
FH	Key
FT	Region
FT	Region
FT	Region
FT	Region
FT	Region
FT	Region
PN	WO200233115-A2.
PD	25-APR-2002.
PF	17-OCT-2001; 2001WO-US032316.
PR	17-OCT-2000; 2000US-0241246P.
PA	(CALY) CALIFORNIA INST OF TECHNOLOGY.
PI	Kumagai A, Dunphy WG;
DR	WPI; 2002-454610/48.
DR	N-PSDB; ABK52611.
PT	Novel Claspin polypeptide specifically interacting with chk1 protein
PT	useful for identifying compound that modulates cell cycle progression and
PT	for treating cell proliferative disorder like neoplasm.
PS	Claim 2; Fig 2; 97pp; English.
CC	The present invention relates to a new substantially pure Claspin
CC	polypeptide that specifically interacts with a Chk1 protein, having SQ/TQ
CC	motifs, an isoelectric point of 4.5 and at least one nuclear localisation

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:37:55 ; Search time 141.414 Seconds
(without alignments)
3793.838 Million cell updates/sec

Title: US-09-982-091A-2
Perfect score: 6604
Sequence: 1 MAALCEEQVLEPEDISLK.....RFRDSTPTVKSRSIFQLLE 1285

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1862994 seqs, 417510619 residues

Total number of hits satisfying chosen parameters: 1862994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6604	100.0	1285	9 US-09-982-091A-2	Sequence 2, Applt
2	2687	40.7	1332	9 US-09-982-091A-4	Sequence 4, Applt
3	784.5	11.9	1812	20 US-11-097-143-858	Sequence 858, App
4	391	5.9	2454	17 US-10-732-923-8884	Sequence 8884, Ap
5	390.5	5.9	6815	20 US-11-097-143-27225	Sequence 27225, A
6	378.5	5.7	3257	20 US-11-097-143-29298	Sequence 29298, A
7	373	5.6	2492	16 US-10-697-526-2	Sequence 2, Applt
8	373	5.6	2492	17 US-10-732-923-8883	Sequence 8883, Ap
9	372	5.6	2375	16 US-10-408-765A-277	Sequence 277, App
10	372	5.6	2375	17 US-10-732-923-8881	Sequence 8881, Ap
11	372	5.6	2492	17 US-10-732-923-8882	Sequence 8882, Ap

12	372	5.6	2492	18 US-10-934-998-76	Sequence 76, Appl
13	371	5.6	2288	17 US-10-732-923-8885	Sequence 8885, Ap
14	371	5.6	2288	18 US-10-934-998-100	Sequence 100, App
15	371	5.6	2337	17 US-10-732-923-8888	Sequence 8888, Ap
16	371	5.6	2375	17 US-10-732-923-8880	Sequence 8880, Ap
17	365.5	5.5	1234	17 US-10-741-849-7132	Sequence 7132, Ap
18	354	5.4	1398	15 US-10-094-466-30	Sequence 30, Appl
19	354	5.4	1400	9 US-09-764-176-7	Sequence 7, Applt
20	341.5	5.2	1909	17 US-10-732-923-3341	Sequence 3341, Ap
21	341	5.2	2897	20 US-11-097-143-2334	Sequence 2334, Ap
22	338.5	5.1	1444	16 US-10-437-963-150334	Sequence 150334,
23	331.5	5.0	3111	20 US-11-097-143-7773	Sequence 7773, Ap
24	329.5	5.0	2476	10 US-10-824-574-7	Sequence 8136, Ap
25	329.5	5.0	2476	17 US-10-732-923-8136	Sequence 7646, Ap
26	324	4.9	1881	14 US-10-032-585-7646	Sequence 3334, Ap
27	323.5	4.9	1875	17 US-10-732-923-3334	Sequence 3329, Ap
28	321.5	4.9	1015	17 US-10-732-923-3329	Sequence 254, App
29	321.5	4.9	3225	16 US-10-408-765A-254	Sequence 22285, A
30	320.5	4.9	1875	15 US-10-369-493-22285	Sequence 3335, Ap
31	320.5	4.9	1875	17 US-10-732-923-3335	Sequence 10224, A
32	318	4.8	1690	20 US-11-097-143-10224	Sequence 10311, A
33	318	4.8	1690	20 US-11-097-143-10311	Sequence 5129, Ap
34	317.5	4.8	2020	15 US-10-369-493-5129	Sequence 3342, Ap
35	317.5	4.8	2020	15 US-10-369-493-3342	Sequence 1586, Ap
36	317	4.8	2910	17 US-10-732-923-3342	Sequence 3330, Ap
37	315.5	4.8	1268	17 US-10-732-923-3332	Sequence 3331, Ap
38	315.5	4.8	1790	15 US-10-369-493-1586	Sequence 31983, A
39	315.5	4.8	1790	17 US-10-732-923-3330	Sequence 31983, A
40	315.5	4.8	1790	17 US-10-732-923-3331	Sequence 31983, A
41	315.5	4.8	2768	20 US-11-097-143-311-4	Sequence 4, Applt
42	314.5	4.8	3899	14 US-10-171-311-8	Sequence 8, Applt
43	314.5	4.8	3917	14 US-10-171-311-8	Sequence 2, Applt
44	310.5	4.7	3907	14 US-10-171-311-2	Sequence 5739, Ap
45	310.5	4.7	3908	18 US-10-756-149-5739	

ALIGNMENTS

RESULT 1
US-09-982-091A-2
; Sequence 2, Application US/09982091A
; Patent No. US20020151030A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: KUMAGAI, Akiko
; APPLICANT: DUNPHY, William
; TITLE OF INVENTION: CLASPIN PROTEINS AND METHODS OF USE THEREOF
; FILE REFERENCE: CIT1320-1
; CURRENT APPLICATION NUMBER: US/09/982,091A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: US 60/241,246
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent version 3.1
; SEQ ID NO 2
; LENGTH: 1285
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-982-091A-2

QY	1	MAALCEEQVLEPEDISLKIVETSDSGQSCCEMADONKLLGCVEDKDTDELIVRKKS	60
DB	1	MAALCEEQVLEPEDISLKIVETSDSGQSCCEMADONKLLGCVEDKDTDELIVRKKS	60
QY	61	KKKEVLVDSDELEMRNFADNVXGHSNDNEENETWSAYREKPKIRSAVLSDNSDHE	120
DB	61	KKKEVLVDSDELEMRNFADNVXGHSNDNEENETWSAYREKPKIRSAVLSDNSDHE	120

Query Match 100.0%; Score 6604; DB 9; Length 1285;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:42:31 ; Search time 156.144 Seconds
(without alignments)
4214.186 Million cell updates/sec

Title: US-09-982-091A-2
Perfect score: 6604
Sequence: 1 MAALCEEQVLEPEDISLK.....RFRDSTPTVKSRSIFQLLE 1285

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6604	100.0	1285	2 Q9DF50	Q9df50 xenopus lae
2	2795	42.3	1315	2 Q80YR7	Q80yr7 mus musculu
3	2687	40.7	1332	2 Q9HAM4	Q9haw4 homo sapien
4	1101	16.7	411	2 Q69GM2	Q69gm2 mus musculu
5	1098.5	16.6	693	2 Q8IWI1	Q8iwl1 homo sapien
6	861.5	13.0	462	2 Q9BX02	Q9bx02 homo sapien
7	786	11.9	1465	2 Q8IRB5	Q8irb5 drosophila
8	722	10.9	289	2 Q9BX01	Q9bx01 homo sapien
9	579.5	8.8	835	2 Q7PR25	Q7pr25 anopheles g
10	496.5	7.5	701	2 Q8MYR6	Q8myr6 drosophila
11	462.5	7.0	725	2 O17838	O17838 caenorhabdi
12	421.5	6.4	3404	2 Q7RJB3	Q7rjb3 plasmodium
13	420	6.4	1464	2 Q871F7	Q871f7 neurospora
14	416.5	6.3	2081	2 Q9LH98	Q9lh98 arabidopsis
15	410	6.2	17903	2 Q9V7L4	Q9v7l4 drosophila
16	409	6.2	7210	2 Q9V7G8	Q9v7g8 drosophila
17	409	6.2	9270	2 Q8MLD9	Q8mld9 drosophila
18	408	6.2	2910	2 Q9FND5	Q9fnd5 arabidopsis
19	404	6.1	1661	2 Q06166	Q06166 plasmodium
20	396	6.0	16215	2 Q9NFS3	Q9nfs3 drosophila
21	393.5	6.0	1607	2 Q6CDM9	Q6cdm9 yarrowia li
22	392	5.9	1510	2 Q25920	Q25920 plasmodium
23	390.5	5.9	18074	2 Q917U4	Q917u4 drosophila
24	390	5.9	2301	2 Q6BG00	Q6bg00 parametium
25	389.5	5.9	17352	2 Q9SYM2	Q9sym2 procambarius
26	389	5.9	791	2 Q9DGL1	Q9dgl1 fugu rubrip
27	386.5	5.9	1434	2 Q8I492	Q8i492 plasmodium
28	380	5.8	5458	2 Q9U459	Q9u459 plasmodium
29	378.5	5.7	3257	2 Q9V736	Q9v736 drosophila
30	375.5	5.7	5507	2 Q8IHN3	Q8ihn3 plasmodium
31	373	5.6	2492	1 ATRX_HUMAN	P46100 homo sapien

32	373	5.6	2762	2 Q9VGW4	Q9vgw4 drosophila
33	370	5.6	2492	1 ATRX_PANTR	Q7ygm4 pan troglod
34	368	5.6	18519	2 Q8ISF6	Q8isf6 caenorhabdi
35	368	5.6	18534	2 Q8ISF7	Q8isf7 caenorhabdi
36	367	5.6	5412	1 FUTS_DROME	Q9w596 drosophila
37	365	5.5	2492	1 ATRX_PONPY	Q7ygm3 pongo pygma
38	364.5	5.5	1481	2 Q8EVB9	Q8evb9 mycoplasma
39	364.5	5.5	2116	1 MYS2_DICDI	Q8ebv9 plasmodium
40	362	5.5	10578	2 Q8ISF5	P08799 dictyosteli
41	359.5	5.4	1110	2 Q91255	Q8isf5 caenorhabdi
42	358	5.4	1871	2 Q9SRD5	Q91255 petromyzon
43	352	5.3	1596	2 Q8IJ44	Q9brd5 arabidopsis
44	351.5	5.3	1430	2 Q9VGZ9	Q8ij44 plasmodium
45	350	5.3	1431	2 Q96T23	Q9vgz9 drosophila
					Q96t23 homo sapien

ALIGNMENTS

RESULT 1

ID	Q9DF50	PRELIMINARY;	PRT;	1285 AA.
AC	Q9DF50;			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Claspin.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_Taxid=8355;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21000493; Pubmed=11090622;			
RA	Kumagai A., Dunphy W.G.;			
RT	"Claspin, a novel protein required for the activation of Chk1 during a			
RT	DNA replication checkpoint response in Xenopus egg extracts.";			
RL	Mol. Cell 6:839-849(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Kumagai A., Dunphy W.G.;			
RL	Submitted (May-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF297867; AAG24516.2; -			
SQ	SEQUENCE 1285 AA; 145981 MW; A302479768FDBA7D CRC64;			

Query Match 100.0%; Score 6604; DB 2; Length 1285;
Best Local Similarity 100.0%; Pred. No. 1.2e-234;
Matches 1285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAALCEEQVLEPEDISLKIVETSDSGGSCEMADQNKLLGVEDKDTDEILVRKKS	60
DB	1	MAALCEEQVLEPEDISLKIVETSDSGGSCEMADQNKLLGVEDKDTDEILVRKKS	60
QY	61	KKKEVLVDSDEELEMRNPAADNVKGHSNDNEENETMSAYREKPRKIRSAVLDSNDHE	120
DB	61	KKKEVLVDSDEELMERNFADNVKGHSNDNEENETMSAYREKPRKIRSAVLDSNDHE	120
QY	121	LDVQISTSQNAAEIPSEHSDLEKETHTVKPKTSKSLKKQDTNKEIIVKNSKRKIPKE	180
DB	121	LDVQISTSQNAAEIPSEHSDLEKETHTVKPKTSKSLKKQDTNKEIIVKNSKRKIPKE	180
QY	181	KIKRRTQKSKAVAEARPNLNDGCLITDGDLPDNGVENEMDSNEEDSLAIRAKMKS	240
DB	181	KIKRRTQKSKAVAEARPNLNDGCLITDGDLPDNGVENEMDSNEEDSLAIRAKMKS	240
QY	241	LNSHSAENFEDFELDTGEGSPKPKRKAARLGEKAMKOMHSETQRLIRBSSVSLPYH	300
DB	241	LNSHSAENFEDFELDTGEGSPKPKRKAARLGEKAMKOMHSETQRLIRBSSVSLPYH	300
QY	301	LPEPKTIHDFEKKRPRPLCOGNAMQLIKSTKQPCTEEEKKKPNEEICAEVPEFDVSKED	360
DB	301	LPEPKTIHDFEKKRPRPLCOGNAMQLIKSTKQPCTEEEKKKPNEEICAEVPEFDVSKED	360

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

```
Run on:      October 25, 2005, 11:53:46 ; Search time 36.3355 Seconds
              (without alignments)
              3402.693 Million cell updates/sec
```

Title: US-09-982-091A-2

Perfect score: 6604

Sequence: 1 MAALCEEQVLEPEDISK.....RFKRDPSTPVKSRSIFQLE 1285

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database : PIR_79:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	462.5	7.0	725	2	T21363	hypothetical protee
2	392	5.9	1526	2	A45605	mature-parasite-in
3	364.5	5.5	2116	2	A26655	myosin heavy chain
4	360	5.5	3488	2	T34418	hypothetical protee
5	359.5	5.4	1110	2	IS1116	NF-180 - sea lamp
6	358	5.4	1871	2	D96796	probable heat shooc
7	355.5	5.4	5327	2	T13564	microtubule-associ
8	341.5	5.2	1909	2	A45592	liver stage antigen
9	340.5	5.2	1658	2	S55101	hypothetical protee
10	336.5	5.1	1898	1	A45973	trichohyalin - hum
11	328.5	5.0	2447	2	T16870	hypothetical protee
12	323	4.9	1624	2	T25592	hypothetical protee
13	322.5	4.9	1922	2	T00637	hypothetical protee
14	322	4.9	1271	2	A45555	glutamate rich pro
15	321.5	4.9	3225	2	IS2300	giantin - human
16	321.5	4.9	3259	1	A56539	giantin - human
17	320.5	4.9	1558	2	B71603	RESA-H3 antigen PP
18	320.5	4.9	1875	2	S38173	myosin-like proteol
19	318.5	4.8	678	2	AS4514	glutamic acid-rich
20	317.5	4.8	2020	2	T21174	hypothetical protee
21	315.5	4.8	1621	2	A82255	hypothetical protee
22	315.5	4.8	1790	2	S67593	transport protein
23	315	4.8	1744	2	JH0720	tanabin - African
24	313	4.7	1979	2	C71622	hypothetical protee
25	312	4.7	2139	2	T18296	myosin heavy chain
26	311.5	4.7	1019	2	T50251	hypothetical colle
27	310	4.7	1233	2	S56271	hypothetical protee
28	309	4.7	1877	2	T21861	hypothetical protee
29	307.5	4.7	2364	2	A56577	microtubule-associ

30	307	4.6	1690	2	T13030	microtubule bindin
31	306	4.6	1939	2	T18372	repeat Organellar
32	305	4.6	2484	2	T26216	hypothetical protee
33	305	4.6	2607	2	T26215	hypothetical protee
34	304	4.6	2464	1	QRMSPI	microtubule-associ
35	303	4.6	5170	2	T15348	hypothetical protee
36	302.5	4.6	1999	1	S21801	myosin heavy chain
37	302.5	4.6	2738	2	E88320	protein F07A11.6 I
38	300	4.5	1620	2	S61535	nucleotide-binding
39	300	4.5	1827	2	T16270	hypothetical protee
40	299.5	4.5	2954	2	T14156	kinesin-related pr
41	298.5	4.5	1992	2	A47297	myosin heavy chain
42	297.5	4.5	2722	2	T20532	hypothetical protee
43	295	4.5	1940	1	S04090	myosin heavy chain
44	294.5	4.5	1927	2	A59236	embryonic muscle m
45	294	4.5	2253	2	T30336	nuclear/mitotic ap

ALIGNMENTS

```

RESULT 1
T21363
hypothetical protein F25H5.5 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T21363
R/Steward, C.
submitted to the EMBL Data Library, October 1996
A/Reference number: Z19412
A/Accession: T21363
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-725 <WIL>
A/Cross-references: UNIPROT:O17838; EMBL:Z81068; PIDN:CAB02986.1; GSPDB:GN00019; CESP:F2
A/Experimental source: clone F25H5
C/Genetics:
A/Gene: CESP:F25H5.5
A/Map position: 1
A/Introns: 5/3; 108/1; 129/1; 233/2; 294/2; 529/2; 601/3

```

Query Match	7.0%;	Score 462.5;	DB 2;	Length 725;
Best Local Similarity	24.8%;	Pred. No. 8.1e-12;		
Matches 210;	Conservative 122;	Mismatches 283;	Indels 233;	Gaps 36;

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 25, 2005, 11:40:50 ; Search time 137.977 Seconds
(without alignments)
3601.964 Million cell updates/sec

Title: US-09-982-091A-2
Perfect score: 6604
Sequence: 1 MAALCEEQVFLPEPDISLK.....RFKRDSPTVKSRSIFQLLE 1285

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1:	Geneseqp19809:*
2:	Geneseqp19908:*
3:	Geneseqp20009:*
4:	Geneseqp20018:*
5:	Geneseqp20028:*
6:	Geneseqp20038:*
7:	Geneseqp2003bs:*
8:	Geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6604	100.0	1285	5	AAU97586	Aau97586 Xenopus C
2	2691	40.7	1332	7	ADC31665	Adc31665 Human nov
3	2687	40.7	1332	5	AAU97587	Aau97587 Human Cla
4	784.5	11.9	1812	4	ABBS8022	Abbs8022 Drosophil
5	473.5	7.2	191	5	ABBS9891	Abbs9891 Human mac
6	390.5	5.9	6815	4	ABBS6811	Abbs6811 Drosophil
7	378.5	5.7	3257	4	ABB67502	Abb67502 Drosophil
8	373	5.6	2492	8	ADQ89534	Adq89534 Human ATR
9	372	5.6	2375	7	ADJ68471	Adj68471 Human hea
10	372	5.6	2492	8	ADK60200	Adk60200 Angiogene
11	372	5.6	2492	8	ADK60501	Adk60501 Angiogene
12	372	5.6	2492	8	ADP73124	Adp73124 Angiogene
13	371	5.6	2288	8	ADK60224	Adk60224 Angiogene
14	371	5.6	2288	8	ADK60525	Adk60525 Angiogene
15	371	5.6	2288	8	ADP73148	Adp73148 Angiogene
16	365.5	5.5	1234	8	ADP98957	Adp98957 C. albica
17	354	5.4	1398	5	ABG97491	Abg97491 Human NOV
18	354	5.4	1400	4	AAB83348	Aab83348 AAP-2 pro
19	353	5.3	1445	6	ABR41365	Abt41365 Human DIT
20	341	5.2	2897	2	ABBS8514	Abbs8514 Drosophil
21	336.5	5.1	1898	2	AAV30795	Aav30795 A human t
22	336.5	5.1	1898	7	ADD48869	Ad448869 Human Pro
23	333.5	5.0	2228	7	ABR61599	Abt61599 Human gol
24	333.5	5.0	2230	6	ABU07445	Abu07445 Protein d
25	333.5	5.0	2230	7	ABR61600	Abt61600 Human gol

26	333.5	5.0	2250	7	ABR61601	Abt61601 Human gol
27	333.5	5.0	2252	7	ABR61602	Abt61602 Human gol
28	331.5	5.0	3111	4	ABBS60327	Abbs60327 Drosophil
29	329.5	5.0	2476	7	ADFI1512	Adfi1512 Murine AT
30	327.5	5.0	1584	8	ABM80133	Abm80133 Tumour-as
31	324	4.9	1881	5	ABP73809	Abp73809 Candida a
32	321.5	4.9	3225	7	ADJ68448	Adj68448 Human hea
33	321.5	4.9	3259	7	ADE56037	Ades6037 Human Pro
34	321.5	4.9	3259	7	ADE56033	Ades6033 Human Pro
35	320.5	4.9	1558	3	AAB18324	Aab18324 Plasmodiu
36	320.5	4.9	1875	6	ABRS53560	Abt53560 Protein s
37	320.5	4.9	1875	7	ADK64380	Adk64380 Disease t
38	320.5	4.9	1875	8	ADS43855	Adt43855 Bacterial
39	318	4.8	1690	4	ABB61144	Abb61144 Drosophil
40	318	4.8	1690	4	ABB61173	Abb61173 Drosophil
41	317.5	4.8	2020	8	ADN22475	Adn22475 Bacterial
42	317.5	4.8	2020	8	ADN22476	Adn22476 Bacterial
43	315.5	4.8	1790	6	ABR53116	Abt53116 Protein s
44	315.5	4.8	1790	7	ADK63040	Adk63040 Disease t
45	315.5	4.8	1790	8	ADN18933	Adn18933 Bacterial

ALIGNMENTS

RESULT 1	
AAU97586	
ID	AAU97586 standard; protein; 1285 AA.
XX	
AC	AAU97586;
XX	
DT	27-AUG-2002 (first entry)
XX	
DE	Xenopus Claspin protein.
XX	
KW	Chk1 protein; SQ/TQ motif; isoelectric point; cell cycle progression;
KW	nuclear localisation signal; DNA replication checkpoint; benign neoplasm;
KW	cell proliferative disorder; malignant neoplasm; frog; claspin.
XX	
OS	Xenopus sp.
XX	
FH	Key
FT	Region
FT	Location/Qualifiers
FT	158..174
FT	/note= "Nuclear localisation signal site"
FT	312..316
FT	/note= "Nuclear localisation signal site"
FT	1078..1084
FT	/note= "Nuclear localisation signal site"
XX	
PN	WO200233115-A2.
XX	
PD	25-APR-2002.
XX	
PF	17-OCT-2001; 2001WO-US032316.
XX	
PR	17-OCT-2000; 2000US-0241246P.
XX	
PA	(CALY) CALIFORNIA INST OF TECHNOLOGY.
PI	Kumagai A, Dumphy WG;
XX	
DR	WPI; 2002-454610/48.
XX	
XX	N-PSDB; ABK52610.
PT	
PT	Novel Claspin polypeptide specifically interacting with chk1 protein
PT	useful for identifying compound that modulates cell cycle progression and
PT	for treating cell proliferative disorder like neoplasm.
XX	
PS	Claim 2; Fig 1; 97pp; English.
XX	
CC	The present invention relates to a new substantially pure Claspin
CC	polypeptide that specifically interacts with a Chk1 protein, having SQ/TQ
CC	motifs, an isoelectric point of 4.5 and at least one nuclear localisation